1

2 MR LUKAS J MUSHER (Orcid ID : 0000-0003-3490-0341)

3 MR PETER GALANTE (Orcid ID : 0000-0002-7025-3551)

4 DR. JERRY HUNTLEY (Orcid ID : 0000-0001-6216-2361)

5

6

7 Article type   : Research Paper

8

9

10 **TITLE:** Shifting ecosystem connectivity during the Pleistocene drove diversification and

11 gene-flow in a species-complex of Neotropical birds (Tityridae: Pachyramphus)

12

13 **RUNNING TITLE:** Diversification with gene-flow in Neotropical becards

14

15 **AUTHORS:**

16 **Lukas J. Musher**[*], Department of Ornithology and The Richard Gilder Graduate

17 School, American Museum of Natural History, Central Park West @ 79th Street, New

18 York, NY 10024; lmusher@amnh.org

19 **Peter J. Galante**, Center for Biodiversity and Conservation, American Museum of

20 Natural History, Central Park West @ 79th Street, New York, NY 10024

21 **Gregory Thom**, Department of Ornithology, American Museum of Natural History,

22 Central Park West @ 79th Street, New York, NY 10024

23 **Jerry W. Huntley**, Department of Ornithology, American Museum of Natural History,

24 Central Park West @ 79th Street, New York, NY 10024

25 **Mary E. Blair,** Center for Biodiversity and Conservation, American Museum of Natural

26 History, Central Park West @ 79th Street, New York, NY 10024

27 [*]Corresponding author

28

29

30

31

32

42

43 **ABSTRACT:**
44 **Aim:** We aim to test the biogeographic drivers of diversification and gene-flow at the
45 Isthmus of Panama using a species complex of suboscine birds as a case study. We
46 specifically evaluate whether diversification in these birds is better explained by
47 continuous parapatry or a Refuge Model of periodic isolation and gene-flow due glacial
48 cycling.
49 **Location:** The Isthmus of Panama (Neotropics)
50 **Taxon:** *Pachyramphus aglaiae* and *P. homochrous* (Aves: Tityridae)
51 **Methods:** We develop an approach to distinguish among the two biogeographic
52 hypotheses –parapatric ecological speciation versus climatically-mediated speciation–
53 by making explicit predictions for demographic history, niche evolution, and change in
54 geographic connectivity over time. We sequenced genome-wide markers
55 (ultraconserved elements) to estimate the evolutionary and demographic history of this
56 group. We applied both phylogenomic network analyses and demographic modeling
57 using a supervised machine learning approach. These genetic analyses were combined
58 with a novel distribution modeling method that estimates the probability of interspecies
59 contact as a function of climatic conditions through time.
60 **Results:** We found that both spatial and genetic analyses revealed concordant results.
61 All speciation events occurred during the Pleistocene and were characterized by non-

62 continuous gene-flow, supporting a scenario of climate-mediated diversification. Spatial

63 connectivity was highest at present, consistent with our best demographic model of

64 secondary contact.

65 **Main Conclusions:** This study exemplifies a mechanism by which speciation,

66 dispersal, and introgression unfold in an important region for Neotropical diversification

67 –the Isthmus of Panama– where periods of *both* isolation and introgression probably

68 drive diversification. Overall, our results are consistent with the Refuge Model of biotic

69 diversification, but suggest that introgression may be a crucial yet underappreciated

70 component of this classic paradigm.

71

72 **KEYWORDS:** refuge hypothesis; statistical phylogeography; neural network; ecological

73 niche modeling; Isthmus of Panama; Pleistocene; introgression

74

75

76

77

78

79 **INTRODUCTION:**

80   Allopatric speciation is typically considered the primary mode of speciation in

81 vertebrates (Coyne and Orr 2004; Provine 2004), but there is growing evidence that

82 gene-flow often occurs during clade diversification (Gompert et al. 2014; Thom et al.

83 2018; Burbrink and Gehara 2018). Although introgression is generally thought to

84 hamper differentiation by homogenizing allele frequencies among populations (Slatkin

85 1987), it can also act as a creative force (Gompert et al. 2006; Norris et al. 2015;

86 Racimo et al. 2015; Barrera-Guzmán et al. 2018). Thus, understanding which processes

87 promote differentiation in the presence of gene-flow is a persistent question in

88 evolutionary biology.

89   One mechanism by which differentiation with gene-flow may occur is through

90 parapatric ecological speciation (Schluter 2001). Ecological speciation occurs when

91 divergent selection due to environmental gradients promotes population differentiation

92 and reproductive isolation despite ongoing gene-flow (Nosil 2012). Here, incipient

93   species adapt to local environments while continuously exchanging alleles, leading to

94   niche evolution (Blair et al. 2013). Yet for a significant portion of species, population

95   differentiation has occurred in the absence of strong environmental gradients (Cadena

96   et al. 2012). Instead, speciation often occurs across biogeographic barriers (i.e., regions

97   of unsuitable habitat) while maintaining some degree of gene-flow through time and

98   space as population connectivity fluctuates with changing environmental conditions (i.e.,

99   isolation-with-migration) (Weir et al. 2015; Winger 2017; He et al. 2019). Under this

100   scenario, barriers to gene-flow are semipermeable spatially (e.g. a river that narrows at

101   the headwaters) and/or temporally (e.g. a river that grows and shrinks through time) to

102   allow for ongoing migration.

103   The semipermeability of barriers can be summarized as two extremes along a

104   continuum. At one extreme, individuals of vagile species may disperse long distances to

105   overcome barriers and hybridize with individuals of populations or species that

106   differentiated in the past (Oswald et al. 2019). This might be particularly expected in

107   archipelagos where the migration of individuals between populations can only be

108   maintained by crossing existing barriers (Weeks and Claramunt 2014). At the other

109   extreme, continental barriers may break down over time, thereby increasing

110   interpopulation migration (Musher et al. 2019; Albert et al. 2017). Thus, populations may

111   experience episodic fragmentation and periodic gene-flow (Zarza et al. 2016; He et al.

112   2019). Under such conditions, the homogenizing effects of gene-flow may be overcome

113   by additional intrinsic mechanisms, such as pre- or post-zygotic isolation, operating at

114   periods of both fragmentation and contact (Provost, Mauck, and Smith 2018; Hoskin et

115   al. 2005).

116   On continents, change in ecosystem connectivity may be a key process affecting

117   patterns of biodiversity (Webb 1991; Kozak and Wiens 2006; Prates et al. 2016; Albert

118   et al. 2017). One frequently explored mechanism that might explain variation in

119   connectivity is historical climate change, which generated pronounced shifts in the

120   distribution of ecosystems and their associated biological communities (Hewitt 2000,

121   2004). During the Quaternary, global climate cycled between warmer/wetter conditions

122   and cooler/dryer conditions, though such changes were not uniformly distributed across

123   the globe (Cheng et al. 2013; Hewitt 2000). These cycles are thought to be closely tied

124   to the fragmentation and connection of ecosystems across the tropics and thus the

125   diversification of tropical biotas (Vuilleumier 1971; Huntley et al. 2018; Potter et al.

126   2018; Musher et al. 2019). When species track their ancestral niches with changing

127   climate (Wiens and Graham 2005; Kozak and Wiens 2006; Tingley et al. 2009),

128   environmental change can affect diversification processes by either reducing population

129   connectivity via habitat fragmentation (Haffer 1969; Mayr and O'Hara 1986; Claramunt

130   and Cracraft 2015) or increasing connectivity via the formation of habitat corridors

131   (Prates et al. 2016; Trujillo-Arias et al. 2018; Flantua et al. 2019).

132   Here we test the influence of climate change and niche evolution in driving

133   diversification and gene-flow across the Isthmus of Panama, where the episodic wax

134   and wane of ecosystem connectivity due to climate change is considered to have

135   played a significant role in diversification (Webb 1991; Vrba 1992; Smith, Amei, and

136   Klicka 2012). To test competing hypotheses, we specifically utilize a species complex of

137   *Pachyramphus* becards that radiated across Central America and northwestern South

138   America. Recent taxonomic work has highlighted a complex of four species across this

139   region whose histories may have included diversification with gene-flow (Musher and

140   Cracraft 2018): *P. aglaiae* (western Mexico), *P. sumichrasti* (eastern Mexico to western

141   Panama), *P. canescens* (eastern Panama to northern Colombia), and *P. homochrous*

142   (southwestern Colombia to northwestern Peru). All taxa are tied to semi-open forested

143   systems, such as gallery and seasonally dry forests, and the edges of lowland

144   evergreen forests (Parker et al. 1996).

145   Although the breeding ranges of *P. sumichrasti* and *P. canescens* are separated

146   at present by more than 400 km around the Isthmus of Panama (Webster 1963), and

147   genetic data support a deep split between these two non-sister species (Musher et al.

148   2019), other patterns suggest the possibility of introgression. For example, phylogeny

149   based on mitochondrial DNA shows a sister relationship between *P. sumichrasti* and *P.*

150   *canescens* plus *P. homochrous*, a relationship that conflicts with the species tree and is

151   potentially the result of ancestral hybridization (Musher and Cracraft 2018). Additionally,

152   males of *P. sumichrasti* are characterized across most of their distribution as having a

153   distinct rose-colored throat, but populations become darker moving southeast, and, like

154   *P. canescens*, lack the distinctive rose-throat in southeastern populations (Webster

155   1963). Furthermore, in *P. sumichrasti* and *P. aglaiae*, phenotypic variation appears to

156   follow Gloger's Rule, as individuals tend to be darker in wetter regions and paler in dryer

157   regions, suggesting that disruptive selection could also play a role within this complex

158   (Webster 1963).

159   Using sequence capture of ultraconserved elements (UCEs), we first test

160   whether diversification in this group is characterized by a history of pure isolation or one

161   in which gene-flow has occurred. Then, using a combination of demographic,

162   phylogenomic, and ecological niche modeling we test whether this history is better

163   explained by parapatric ecological or climate-mediated speciation-with-gene-flow.

164

165   *Predictions for competing hypotheses*

166   Here, we develop a three-pronged approach to distinguish among biogeographic

167   hypotheses that evaluates each hypothesis based on predictions for demographic

168   history, niche evolution, and change in geographic connectivity over time. We

169   specifically test two modes of speciation. First, we test parapatric ecological speciation

170   (Nosil 2012), which predicts continuous gene-flow, niche divergence, and constant

171   geographic connectivity over time. Such a model would be consistent with adaptation to

172   local environmental conditions that overwhelms the homogenizing effects of gene-flow.

173   Second, we test climate-mediated speciation, wherein differentiating populations

174   experience gene-flow that becomes reduced during times of decreased ecosystem

175   connectivity. This model predicts non-continuous gene-flow, niche conservatism, and

176   geographic connectivity that varies with climate change. Although gene-flow is not

177   required for ecological speciation in the broad sense (Nosil 2012), herein we use the

178   term "ecological speciation" to specifically mean *parapatric* ecological speciation.

179

180   **METHODS:**

181

182   *Sequencing and bioinformatics*

183   We extracted genomic DNA from a total of 21 specimens across 19 localities

184   (Table S1; Figure 1). These included 17 vouchered fresh tissues and four toe pads from

185   historical museum study skins. To sequence UCEs, DNA extracts were sent to

186 RapidGenomics® (Gainesville, FL) for DNA shearing, library preparation, and Illumina

187 sequencing. We processed raw reads from Illumina output in two ways. First, we used

188 the Phyluce pipeline (Faircloth 2016) to trim, clean, and assemble reads. Second, we

189 used a custom pipeline to map reads to assembled UCE loci, phase all loci for all

190 individuals, and call single nucleotide polymorphisms (SNPs). Details of DNA extraction

191 and bioinformatic protocols can be found in the Electronic Supplementary Material

192 (ESM).

193

194 *Population and species delimitation*

195 To identify genetically distinct populations without *a priori* bias, we used

196 STRUCTURE v2.3.4 (Hubisz et al. 2009). We sampled one random SNP from each

197 locus in our 75% complete dataset (loci represented by a minimum of 75% of

198 individuals), and performed 20 runs of 5×10^5 generations after a burnin 10^5

199 generations for each of four K-values (K=2–5). We then used CLUMPP (Jakobsson and

200 Rosenberg 2007) to optimally align runs by permuting 10^5 replicates for each K value.

201 We then found the optimal K by choosing the model with the highest ad hoc ΔK value

202 (Evanno, Regnaut, and Goudet 2005). We additionally present population assignment

203 results for all four K-values to better characterize the variance in population structure

204 represented in the data.

205 Many species concepts define species as populations or sets of populations that

206 are reproductively cohesive over evolutionary time (De Queiroz 2007). To test for

207 species boundaries in this group, we used Bayesian Phylogenetics and Phylogeography

208 (BP&P) v3.4 (Rannala and Yang 2013), which statistically evaluates lineage

209 independence using a multi-species coalescent model on sequence data. Because

210 previous work has shown that the species tree relationships are robust based on results

211 from ASTRAL (Mirarab et al. 2014; Musher and Cracraft 2018), we used algorithm A10,

212 which implements a reversible-jump Markov chain Monte Carlo to move between

213 different species delimitation models within a fixed guide tree. Because the population

214 assignments based on STRUCTURE were imbedded within our BP&P models, which

215 were based on previous taxonomic work rather than population assignments. We tested

216   multiple combinations of prior distributions for root divergence time and ancestral

217   mutation-scaled effective population size (Leaché and Fujita 2010) (See ESM).

218

219   *Phylogenomic networks*

220   To characterize the history of diversification and gene-flow, we used two

221   methods. First, we used PhyloNet 3.8.0 (Yu and Nakhleh 2015), which infers the

222   maximum pseudolikelihood phylogenetic network using a set of rooted gene trees while

223   accounting for both incomplete lineage sorting and hybridization. In PhyloNet, reticulate

224   edges are assigned inheritance probabilities, which are related to the proportion of

225   incongruent gene trees that can be explained by hybridization. We generated gene

226   trees using RAxML 8.2.12 (Stamatakis 2014) on unphased sequences recovered in the

227   Phyluce pipeline (75% and 95% complete datasets). Then, we assigned individuals

228   based on BP&P results (plus one outgroup) and applied the *InferNetwork_MPL*

229   algorithm with a single reticulation and 50 pseudolikelihood searches. We repeated this

230   run 20 times and kept the maximum pseudolikelihood topology across all runs.

231   Second, we used TreeMix v1.13 to estimate the maximum likelihood population

232   graph in the presence of gene-flow (Pickrell and Pritchard 2012). TreeMix uses full

233   likelihood to add admixture events to a tree of bifurcating populations. These admixture

234   events or migration "edges" describe populations with ancestry derived from another

235   population. The inferred migration weight of these events is related to the proportion of

236   alleles in a population derived from migration. We assigned individuals to populations

237   based on BP&P results and ran TreeMix using the 75% and 95% complete SNP

238   datasets. Each run was replicated for both m=1 and m=2 for the number of admixture

239   events.

240

241   *Historical demographic inference*

242   To infer the history of isolation and gene-flow between *P. sumichrasti* and *P.*

243   *canescens*, we employed a machine learning approach to fit simulated models to our

244   observed dataset. Sequence data were simulated under several demographic models in

245   PipeMaster (Gehara et al. 2017; Gehara, Mazzochinni, and Burbrink 2019). PipeMaster

246   uses msABC (Hudson 2002; Pavlidis, Laurent, and Stephan 2010) to simulate

247   nucleotide sequences with the same length and number of individuals as loci in the

248   observed dataset. It then uses PopGenome (Pfeifer et al. 2014) to calculate summary

249   statistics for simulated and observed data. Parameters for each simulation were

250   sampled from uniform priors, including a mutation rate for each gene. We simulated

251   data under four scenarios following the assumptions of the aforementioned hypotheses

252   (Figure S1): (M1) isolation without migration, (M2) isolation with continuous migration,

253   (M3) isolation with ancestral migration (migration before the last interglacial, 130kya),

254   (M4) isolation with secondary contact (migration after the last interglacial, <130kya). We

255   used the same set of wide priors for effective population size (min:50,000

256   max:1,500,000), divergence times (min:130,000 max:1,500,000 generations) and

257   migration rate (min:0.1, max: 2.0 migrants per generation, when allowed) for all models.

258   Time priors assume a generation time of 1 year. Given our hypotheses, our primary

259   goal in testing these demographic scenarios was to assess the pattern of historical

260   connection among populations. Although ideally we would evaluate more parameters-

261   rich models (e.g., to incorporate population size change), given our reduced dataset we

262   decided to conservatively test scenarios that exclusively varied in gene-flow. Despite

263   the necessary simplicity of the tested scenarios, they offer a guide to the histories of

264   gene-flow and isolation in these taxa, which should provide the means to distinguish

265   between our hypotheses.

266   The high information content associated with thousands of markers enables high

267   accuracy in model selection even in approaches with a relatively low number of

268   simulations (Gehara, Mazzochinni, and Burbrink 2019). For each model, we performed

269   24,000 simulations and calculated nine population genetic statistics, summarized as

270   means and variances across all loci. We then transformed the summary statistics using

271   principal components analysis (PCA), which was used to evaluate the fit of observed

272   data to simulated models. We explored model selection accuracy using supervised

273   machine learning, running 100 cross-validation replicates of pseudo-observed datasets

274   with the nonlinear neural network algorithm implemented in the ABC R package

275   (Csilléry, François, and Blum 2012). Initial data exploration suggested higher model

276   selection accuracy when retaining a subset of simulations closer to the observed data.

277   Hence, to explore how model selection and parameter estimation could be affected by

278   different numbers of retained simulations we implemented three tolerance values (0.1,

279   0.05 and 0.01, retaining 9,600, 4800 and 960 simulations respectively). To assess

280   model selection accuracy, we calculated the number of times the true model was

281   chosen (Posterior Probability; PP). Additional detail is found in the ESM.

282

283   *Mitochondrial genome (mtGenome) divergence dating*

284   A previous study found that relationships based on mitochondrial DNA conflicted

285   with the species tree suggesting either incomplete lineage sorting or introgression

286   caused incongruence between gene histories (Musher and Cracraft 2018). The node of

287   interest, however, was weakly supported. To improve the accuracy of mitochondrial

288   phylogeny and to independently estimate the timeframe of diversification, we recovered

289   whole mtGenomes from off-target reads as byproduct of UCE sequencing. To estimate

290   the timing and historical relationships of mtDNA divergences, we used BEAST v.2.5.1

291   (Bouckaert et al. 2014) (See ESM). To estimate divergence times, we applied clock rate

292   means and standard deviations to all genes based on estimated rates from a previous

293   study (clock means: ND1=0.025, ND2=0.029, ND3=0.024, ND4=0.022, ND5=0.021,

294   COI=0.016, COII=0.019, COIII=0.019, ATP6=0.026, ATP8=0.019, and cyt *b*=0.014)

295   (Lerner et al. 2011).

296

297   *Ecological niche modeling*

298   To test spatial and ecological predictions, we built ecological niche models

299   (ENMs) for the two non-sister species that meet at the Isthmus of Panama. Because the

300   Isthmus is the location of both historical speciation and potential gene-flow, we only

301   constructed ENMs for *P. sumichrasti* and *P. canescens.* We downloaded locality data

302   from the Global Biodiversity Information Facility (GBIF) for both species. We then

303   filtered and thinned these data to create ENMs through a model tuning exercise. In this

304   exercise, we varied a range of model complexities using regularization multipliers (RMs)

305   and feature classes (FCs) for Maxent v3.4.1 (Phillips, Anderson, and Schapire 2006)

306   using *ENMeval* (Muscarella et al. 2014) (See ESM).

307   To test if suitable habitat has shifted with changing climate, we used the optimal

308   settings from the above tuning exercise to build models using all occurrences. We

309   projected these models to four time periods under the Community Climate System

310   General Circulation Model (CCSM4) (Kiehl and Gent 2004): current (CUR), Mid-

311   Holocene (MH), last glacial maximum (LGM), and last interglacial (LIG) (Otto-Bliesner et

312   al. 2006). Here, models were thresholded using the equal training sensitivity and

313   specificity threshold for the cloglog projection (Phillips et al. 2017). Using this threshold,

314   we compared models for overlap using Schoener's D-value (Schoener 1968), percent

315   concordance (proportion of total geographic distribution shared by both species), and

316   the proportion of the study area predicted as present.

317   To statistically test if the degree of geographic connectivity varied through time,

318   we developed a bootstrapping protocol that randomly samples 'neighborhoods' of fixed

319   radius and quantifies the mean joint probability of occurrence for each neighborhood

320   (henceforth joint suitability). Joint probability was determined by taking the product of

321   the probabilities (projected in raw format) for each species at a given grid cell, and the

322   mean was calculated given all cells in each neighborhood. We applied 1000 replicates

323   to each time slice, which generated a frequency distribution of mean joint suitabilities for

324   each time period. We use this joint suitability measure as a proxy for the degree of

325   connectivity at each time period. To evaluate the effect of relatively large and small

326   neighborhoods we applied this method using neighborhoods of both 50 km and 250 km

327   radii.

328   We then determined whether joint suitability varied with climate by applying a

329   Bayesian analysis of variance in the R package *BayesFactor* (Morey, Rouder, and Jamil

330   2015). We tested alternative models of connectivity change by, first, evaluating whether

331   there was more support for a model where connectivity varied across all time periods

332   versus remaining constant, and second by evaluating the support for connectivity

333   change between each consecutive time period independently. The latter test was

334   evaluated at each time period by taking the ratio of the Bayes factor where all times

335   vary in connectivity to the Bayes factor where two consecutive periods are equivalent in

336   connectivity. Five models were tested in total: LIG = LGM, LGM = MH, MH = current,

337   and LIG ≠ LGM ≠ MH ≠ current, which were all tested against a null model where all

338   time periods were equivalent. We additionally evaluated the PP of change between

339   each time period by sampling the posterior of the model where all periods vary using

340 100,000 MCMC replicates, and quantifying the proportion of samples that were

341 consistent with change between time periods.

342   To test if niches have been phylogenetically conserved, we compared the niche

343 models for similarity and equivalency using *ecospat* (Di Cola et al. 2017). We tested for

344 'niche similarity' by randomly shifting the niches of each species throughout the study

345 area, and then 'niche equivalency' by randomly selecting occurrences from both species

346 to create new datasets and ENMs. Both analyses resulted in a frequency distribution of

347 Schoener's D-values against which we tested against observed values. Niches were

348 determined to be similar or equivalent if the observed D-value was greater than

349 expected under the 1000 randomizations (see ESM).

350

351 **RESULTS:**

352

353 *Population discovery and species delimitation*

354   The optimal K-value was 3 based on STRUCTURE (ΔK = 532.744; Table S2),

355 but assignments for all K-values showed widespread admixture (Figure 2A). Notably,

356 several individuals from *P. sumichrasti* were admixed with respect to the *P. aglaiae*

357 population. These results were consistent with *P. aglaiae* and *P. sumichrasti*

358 representing distinct population clusters, though *P. canescens* and *P. homochrous* were

359 lumped into a single population. BP&P recovered four lineages for all three runs of

360 alternate priors (PP=1.0 for all nodes).

361

362 *Phylogenomic networks*

363   Phylogenetic network results were indicative of a history of gene-flow during

364 diversification of the *Pachyramphus* species complex. In PhyloNet, both 75% (lnL = -

365 535667.93) and 95% (lnL = -421640.77) complete datasets recovered the same

366 topology and inheritance probabilities (Figure 2B). TreeMix also recovered similar

367 results for both SNP datasets, inferring strong migration from *P. sumichrasti* to *P.*

368 *canescens* (75% m1: lnL = 68.89; 75% m2: lnL = 68.91; 95% m1: lnL = 66.07; 95%

369 m2: lnL = 66.93). In m=2 models, TreeMix inferred a second admixture edge from *P.*

370 *aglaiae* to *P. homochrous*, though the migration weight was low for the 75% SNP

371 dataset (Figure 2B).

372

373 *Historical demographic inference*

374   The four models were accurately identified on 82% of the cross-validations,

375 across the tolerance levels applied (Figure S2). The mean PP for each model, obtained

376 with the cross-validations, suggest that our approach is efficient at distinguishing among

377 models (average PP = 0.73 across models and tolerances). Distinct tolerances

378 produced similar model classifications (mean correct classification of 82.5, 80.25 and

379 83.25 for the 0.1, 0.05 and 0.01 tolerances, respectively). The cross-validations for

380 parameter estimation suggest that current effective population sizes and divergence

381 times can be correctly estimated, with high and significant correlation values ($R^2 > 0.9$; p

382 < 0.01) between pseudo-observed datasets and simulations (Figure S3). The lower

383 values for the Pearson's correlation index between simulated and pseudo-observed

384 data were observed for migration parameters ($R^2 \sim 0.5$; p < 0.01), suggesting a reduced

385 capacity of our approach to estimate the magnitude of gene-flow between the two

386 populations, despite high accuracy in distinguishing among models with gene-flow.

387   The PCA suggested that the simulations fit the data well, with the first three

388 components explaining 80% of the variance (Figure S4). At a tolerance of 0.1, the

389 model selection approach supported secondary contact (M4) with relatively high PP,

390 and found that gene-flow occurred from the last interglacial (130ky) towards the present

391 (PP = 0.81; Bayes factor > 6.0; Table 1). However, the PP for the secondary contact

392 model decreased with lower tolerances, and was only marginally better than the

393 continuous migration model (Bayes factor ~ 2; Table 1). Parameter estimation for the

394 secondary contact model supported a larger effective population size for *P. sumichrasti*

395 than *P. canescens*, a divergence time during the early-mid Pleistocene (mean =

396 580,740 generations ago; 95% CI = 407,948 - 865,711 generations ago), and migration

397 rates that overall reflected the priors (Table 2).

398

399 *Mitochondrial divergence dating*

400   We recovered a mtGenome topology that conflicted with the species tree,

401   confirming that either incomplete lineage sorting or ancestral introgression has resulted

402   in discordance between mitochondrial and species histories (Figure 2A). The initial split

403   occurred roughly 710 (820–610) kya, followed by a second split roughly 630 (730–530)

404   kya. The divergence *P. canescens* and *P. homochrous* occurred about 466 (560–380)

405   kya. All nodes were recovered with strong support (PP>0.95). Divergence of this group

406   from its outgroup (not shown) occurred 1,054 (1,303–1,778) kya.

407

408   *Ecological niche modeling*

409   The optimal model for *P. canescens* predicts coastal regions of northern South

410   America and the Isthmus of Panama (Figure 3), in addition to low-to-middle elevations

411   of the northern Andes. The optimal model selected for *P. sumichrasti* showed strong

412   climatic suitability along the Pacific coast of Mexico and Central America as well as high

413   suitability through the Isthmus of Panama (Figure 3).

414   Despite some obvious geographic differences in the predicted distributions of

415   these species, the niches of *P. canescens* and *P. sumichrasti* were conserved (Figure

416   S5). The climatic suitabilities for the two species were more similar (equivalent) than

417   randomly combining occurrences from each species (p = 0.001). We also found that

418   distributional overlap between the two species was higher than expected by chance (p =

419   0.047; Figure S5).

420   Projections of optimal models through time indicate high levels of ecosystem

421   connectivity for both species throughout the Isthmus of Panama. Current and LIG time

422   periods showed the highest levels of niche overlap observed, with drops in overlap at

423   the LGM. *Pachyramphus sumichrasti* consistently showed the highest proportion of the

424   study region predicted as present through each time period. *Pachyramphus canescens*

425   predictions were reduced during the LGM through the Isthmus of Panama (Figure 3).

426   Bootstraps supported a scenario of varying connectivity through time, including

427   for neighborhoods of 50km (Bayes factor > 10^{160}) and 250km radii (Bayes factor >

428   10^{501}) (Figure 3). Specifically, connectivity declined between LIG and LGM (50km Bayes

429   factor > 10^4, PP = 1.00; 250km Bayes factor > 10^{15}, PP = 1.00), increased between

430   LGM and MH (50km Bayes factor > 10^{15}, PP = 1.00; 250km Bayes factor > 10^{49}, PP =

431   1.00), and increased again from MH to current climates (50km Bayes factor > 10^{68}, PP

432   = 1.00; 250km Bayes factor > 10^{251}, PP = 1.00), and that the order (current > LIG > MH

433   > LGM) was strongly supported under both 250km (Bayes factor = 24.00, PP = 1.00)

434   and 50km (Bayes factor = 24.00, PP = 1.00) neighborhoods. Interestingly, many

435   outlying high values were detected by 50km radius bootstrap during LGM climates,

436   perhaps reflecting small patches of joint suitability not detected by larger

437   neighborhoods.

438

439   **DISCUSSION:**

440   Our spatial and genetic results were concordant and supported a scenario of

441   climate-mediated diversification with periodic gene-flow. Phylogenomic and

442   demographic analyses both indicated a history of gene-flow during diversification.

443   Geographic connectivity between *P. sumichrasti* and *P. canescens* varied significantly

444   through time with highest joint suitability during contemporary and lowest during the

445   LGM climates (Figure 3), a scenario that was consistent with our best-fitting

446   demographic model of secondary contact. Accordingly, because gene-flow was not

447   continuous, niches were not divergent, and joint climatic suitability was not constant

448   through time, we reject parapatric ecological speciation in favor of shifting ecosystem

449   connectivity due to climate change. Speciation in *Pachyramphus* is therefore attributed

450   to initial isolation followed by introgression during periods of increased ecosystem

451   connectivity.

452   We also found evidence of admixture between *P. aglaiae* and *P. sumichrasti*, two

453   recently diverged sister taxa, based on Structure (Figure 2). This is consistent with

454   previous work, which found that *P. sumichrasti* was paraphyletic relative to *P. aglaiae*

455   based on a concatenated UCE phylogeny (Musher and Cracraft 2018). Furthermore,

456   TreeMix inferred an unlikely scenario of gene-flow between *P. aglaiae* and *P.*

457   *homochrous,* two non-sister taxa that are geographically non-adjacent. We cannot

458   easily explain this finding, except that the 75% complete dataset (which contains more

459   loci) found this event to be very weak, suggesting that many loci conflicting with this

460   scenario were filtered out to generate the 95% complete dataset. The likelihoods were

461   also very similar between m=1 and m=2 models, suggesting a model with fewer edges

462   would be favored. Still, because both m=1 and m=2 TreeMix models detect admixture

463   between *P. sumichrasti* and *P. canescens*, we argue that this admixture event is better

464   supported.

465

466   *Implications for a Refuge Model of diversification*

467   Our results are consistent with the Refuge Model of biotic diversification, which

468   posits that global cycles of climate warming and cooling during the Quaternary caused

469   expansion and retraction of forested and non-forested biomes (Haffer 1969; Webb

470   1991). However, our data support an alternative view of the Refuge Model in which

471   environmental change has been subtler than the periodic replacement of forest with

472   open savannas (Smith, Amei, and Klicka 2012; Arruda et al. 2018). Recent work on

473   South America has shown that during cooler arid periods, there was likely an increase

474   in tropical deciduous and open ombrophilous forests, in addition to downward

475   elevational shifts of montane ecosystems (Robinson, Brawn, and Robinson 2000;

476   Arruda et al. 2018). Rather than driving pure allopatric isolation, this process would

477   have altered landscape connectivity at ecological scales and in turn varied the

478   probabilities of isolation, local extinction, and dispersal over time. This was evidenced in

479   our ENMs, in which ecological connectivity never ceased; rather, the probability of

480   connectivity varied predictably through time, with the last glacial maximum showing the

481   lowest levels of connectivity, as expected *a priori* (Figure 3).

482   Thus, a model of forest retraction and coalescence may not effectively describe

483   historical dispersal and speciation at the Isthmus of Panama. Along the Isthmus, which

484   already limits dispersal due to a spatial bottleneck (Robinson, Brawn, and Robinson

485   2000), changes in climate may specifically impact variation in rates of dispersal or gene-

486   flow through time. In this way, connectivity between ecosystems across the Isthmus of

487   Panama may be 'flickering' (Flantua and Hooghiemstra 2018). This 'flickering

488   connectivity' may not only explain patterns across the Isthmus of Panama, but also in

489   other regions where connectivity occurs via narrow or 'stepping-stone' corridors, such

490   as between Amazonian and Atlantic Forests of South America, along tropical mountain

491   ranges, or across the Dahomey Gap of western Africa (Mayr and O'Hara 1986; Rull and

492   Nogué 2007; Prates et al. 2016; Trujillo-Arias et al. 2018; Huntley et al. 2019). We

493 therefore suggest that the climatically-mediated scenario of periodic gene-flow that we

494 propose at the Isthmus of Panama could also be an important phenomenon in other

495 regions (Ehrendorfer 1959; Rattenbury 1962; Zarza et al. 2016; He et al. 2019).

496   Moreover, we argue that periodic gene-flow may be a more important component

497 of the Refuge Model than previously realized. For example, cycling between allopatry

498 and sympatry might promote greater net-diversification than either sympatry or allopatry

499 alone (Aguilée, Claessen, and Lambert 2013), and there is a growing body of evidence

500 that suggests gene-flow may promote speciation via adaptive introgression (Delmore et

501 al. 2015; Racimo et al. 2015; Irwin et al. 2018; Marques, Meier, and Seehausen 2019)

502 or hybrid speciation (Barrera-Guzmán et al. 2018). Taken together, these data suggest

503 that introgression, in addition to isolation, is a crucial process affecting diversification.

504   Because real-world data often do not fit neatly into discrete geographic modes of

505 speciation (Schilthuizen 2000; Fitzpatrick, Fordyce, and Gavrilets 2008; Harrison 2012),

506 focusing entirely on cycles of allopatry has rendered the Refuge Model incomplete.

507 Importantly, although we reject a model of parapatric ecological speciation, elements of

508 such a model are probably relevant to the Refuge Model framework, and future studies

509 evaluating the effects of climate on diversification should consider the role of periodic

510 introgression, rather than pure isolation. Though our best-fitting demographic model

511 points to a scenario where gene-flow occurs only between non-sister taxa, a model of

512 continuous migration is second-best for two of three tolerance thresholds and PhyloNet

513 recovered deeper reticulation, suggesting that some historical gene-flow has likely

514 occurred (Table 1).

515

516 *Limitations of our approach*

517   Despite consistency among our results, we draw interpretations with an important

518 caveat. From a phylogeographical perspective, modeling demographic histories is still

519 extremely challenging, even for whole-genome data (Beichman, Phung, and Lohmueller

520 2017). Pulses of isolation and gene-flow between *Pachyramphus* taxa could mislead the

521 necessarily simplified models that we explored. For example, a scenario with multiple

522 pulses of connection could mimic a model of continuous migration. Similarly, long

523 periods of isolation could overwrite signals of historical gene-flow. However, our

524   approach points to a scenario where gene-flow began during the last glacial cycle,

525   despite not having enough resolution to distinguish between multiple pulses of contact

526   or isolation that potentially occurred within this period.

527   In other words, given our limited sampling, we lack the ability to infer population-

528   level processes with the precision needed to look within a single glacial cycle as was

529   done using spatial analysis. Thus, we suggest that our demographic models be

530   interpreted as a rough guide to the processes occurring rather than the true

531   demographic history, the latter of which is undoubtedly more complex. Given this

532   constraint, we were still able to distinguish among continuous and non-continuous

533   models of gene-flow, which were ultimately consistent with the spatial results as neither

534   geographic nor genetic connectivity was continuous through time. Importantly, we found

535   that despite limited sampling, competing demographic models were nonetheless

536   distinguishable within our machine learning framework, suggesting that this approach is

537   useful even for studies with smaller sample sizes (Gehara, Mazzochinni, and Burbrink

538   2019) (see also ESM; Figure S6).

539

540   *Conclusions*

541   In *Pachyramphus*, we provided evidence that varying connectivity across the

542   Isthmus of Panama has led to both differentiation and gene-flow. Reduced connectivity

543   apparently drove repeated speciation as multiple taxa have accumulated over time,

544   while increased connectivity later promoted gene-flow between non-sister taxa (*P.*

545   *canescens* and *P. sumichrasti*), indicating a reticulate history for this group. The timing

546   of initial isolation roughly 400 to 700 kya is probably associated with a major cooling

547   event, as several of the coldest periods in Earth's history occurred over the past 800

548   thousand years (Hewitt 2000).

549   Because climate has varied dramatically over this period, degrees of connectivity

550   and fragmentation between many biomes, globally, have varied as well. Such variation

551   in connectivity during the Pleistocene could also explain a range of macroevolutionary

552   patterns. For example, variable connectivity during the Pleistocene may explain why

553   community-wide patterns of gene-flow are congruent between some Andean dry forests

554 but divergence times are not, a pattern that is perhaps driven by recolonization of locally

555 extinct populations facilitated by cycling connectivity (Oswald et al. 2017).

556   Therefore, our results have important implications for future studies evaluating

557 rates of dispersal or gene-flow through time. Biotic dispersal is often modeled as a

558 stochastic, time-dependent process (Sanmartín, Van Der Mark, and Ronquist 2008),

559 which certainly captures much of the variance in dispersal rates through time and

560 across lineages. Still, rates of within-lineage dispersal are often not evenly distributed

561 over time but are instead influenced by additional factors that affect ecosystem

562 connectivity, which vary the probability of successful dispersal over time (Musher et al.

563 2019; Prates et al. 2016).

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585 **TABLES:**

586

587 **Table 1:** Summary of the neural network model selection approach for three tolerance

588 levels. *M1* - Pure isolation, *M2* - isolation with continuous migration, *M3* - isolation with

589 ancestral migration, *M4* - isolation with secondary contact.

Tolerance	Support	M1	M2	M3	M4
0.1	Posterior probability	0.1218	0.0641	0.0025	**0.8116**
	Bayes factor	6.6612	12.6554	330.4801	**1**
0.05	Posterior probability	0.0001	0.2976	0	**0.7024**
	Bayes factor	10946.532	2.3605	112920.029	**1**
0.01	Posterior probability	0.0034	0.3	0.0415	**0.6551**
	Bayes factor	191.6658	2.1836	15.7961	**1**

590
591
592
593
594
595
596
597

598

599

600 **Table 2:** Summary of the neural network parameter estimation approach for the best

601 model (*M4*) for three tolerance levels. *Ne* - effective population size, *mig* - migration

602 (arrow represents the direction of gene-flow in forward time), *sd* - standard deviation, *µ* -

603 mutation rate.

	Parameter	Ne *P. sum.*	Ne *P. can.*	Divergence time	mig can. -> sum	mig sum. -> can	Mean µ	Sd µ
	Prior	50k – 1500k	50k – 1500k	130k – 1500k	0.5 – 2.0	0.5 – 2.0	1E-11 – 1E-9	1E-11 – 1E-9
tolerance 0.1	**2.5%CI**	169176	41320	407948	0.65	0.81	1.71E-10	3.30E-11
	Median:	931986	238041	549487	1.16	1.55	4.56E-10	3.82E-10
	Mean:	920580	311597	580740	1.16	1.48	4.55E-10	4.37E-10
	Mode:	757072	155208	459940	0.85	1.8	3.70E-10	5.50E-11
	97.5%CI	1582967	788251	865711	1.67	1.89	1.38E-09	1.00E-09
tolerance 0.05	**2.5%CI**	331463	160891	261296	0.51	0.76	1.88E-11	6.66E-11
	Median:	833531	231644	503520	1.25	1.66	4.02E-10	3.54E-10
	Mean:	826280	257834	542706	1.24	1.57	4.04E-10	3.88E-10

	Mode:	718765	204415	387092	0.8	1.88	9.67E-11	7.34E-11
	97.5%CI	1242465	437602	954959	1.97	2.04	9.82E-10	9.93E-10
tolerance 0.01	**2.5%CI**	466803	-269903	-259312	0.39	0.88	4.59E-11	9.92E-11
	Median:	940044	114191	344292	1.17	1.68	3.80E-10	3.89E-10
	Mean:	903653	182340	386831	1.2	1.59	4.14E-10	4.25E-10
	Mode:	1079216	-9023	646281	0.76	1.81	1.54E-10	2.10E-10
	97.5%CI	1210353	950732	1134811	2.03	1.98	1.15E-09	8.70E-10

604

605

606

607 **FIGURES:**

608 **Figure 1:** Map of study region. Red box within inset demarcates the boundaries of the

609 study region. Distributions (dashed lines) and localities (filled circles) of samples are

610 colored by putative taxonomic limits for *P. aglaiae* (pink), *P. sumichrasti* (orange), *P.*

611 *canescens* (green), and *P. homochrous* (blue) are based on results from a previous

612 study (Musher and Cracraft 2018).

613 **Figure 2:** A) Left: a phylogeny inferred in BEAST using the mtGenomes. Posterior

614 Probabilities are shown at each node and red bars represent 95% highest posterior

615 densities. Center: results from STRUCTURE for each of four values of K. Numbers in

616 parentheses correspond to localities in Figure 1. Toe pad specimens are indicated with

617 red text. Right: the species tree topology from a previous study (Musher and Cracraft

618 2018), which was used as a guide tree for BP&P analyses. Posterior probability was 1.0

619 for all speciation events on this tree for all three prior sets used. B) Results from

620 PhyloNet and TreeMix for the 75% (left column) and 95% (right column) complete

621 datasets. Red lines indicate admixture, which are labeled with their inferred inheritance

622 probabilities (PhyloNet) and migration weights (TreeMix). The thickness of the arrows

623 for TreeMix plots are proportional to migration weight.

624 **Figure 3:** The left panel shows the optimal ENM predictions for each species under

625 current climate. The center panel shows models at the Isthmus of Panama for *P.*

626 *sumichrasti* (yellow), *P. canescens* (red), and model overlap (orange). Values for

627 Schoener's D, percent concordance, and proportion of the study region predicted as

628 present (PPP) are shown for each time slice: current (CUR), Mid-Holocene (MH), last

629 glacial maximum (LGM), and last interglacial (LIG) time periods. The right panel shows

630 subsampled joint suitability through time for 50km- (light blue) and 250km- (tawny)

631 radius neighborhoods.

632

633 **DATA AVAILABILITY**: Scripts for replicating analyses are available at

634 https://github.com/lukemusher/speciation-with-gene-flow-at-Isthmus-of-Panama.

635 Cleaned (adapters and low-quality bases removed) are available on dryad at

636 https://doi.org/10.5061/dryad.vt4b8gtp5.

637

638 **LITERATURE CITED**

639 Aguilée, Robin, David Claessen, and Amaury Lambert. 2013. "Adaptive Radiation

640 Driven by the Interplay of Eco-Evolutionary and Landscape Dynamics." *Evolution*

641 67 (5): 1291–1306.

642 Albert, James S., Donald R. Schoolmaster Jr, Victor Tagliacollo, and Scott M. Duke-

643 Sylvester. 2017. "Barrier Displacement on a Neutral Landscape: Toward a Theory

644 of Continental Biogeography." *Systematic Biology* 66 (2): 167–82.

645 Arruda, Daniel M., Carlos E. G. R. Schaefer, Rúbia S. Fonseca, Ricardo R. C. Solar,

646 and Elpídio I. Fernandes-Filho. 2018. "Vegetation Cover of Brazil in the Last 21 Ka:

647 New Insights into the Amazonian Refugia and Pleistocenic Arc Hypotheses." *Global*

648 *Ecology and Biogeography: A Journal of Macroecology* 27 (1): 47–56.

649 Barrera-Guzmán, Alfredo O., Alexandre Aleixo, Matthew D. Shawkey, and Jason T.

650 Weir. 2018. "Hybrid Speciation Leads to Novel Male Secondary Sexual

651 Ornamentation of an Amazonian Bird." *Proceedings of the National Academy of*

652 *Sciences*. https://doi.org/10.1073/pnas.1717319115.

653 Beichman, Annabel C., Tanya N. Phung, and Kirk E. Lohmueller. 2017. "Comparison of

654 Single Genome and Allele Frequency Data Reveals Discordant Demographic

655 Histories." *G3* 7 (11): 3605–20.

656 Blair, M. E., E. J. Sterling, M. Dusch, C. J. Raxworthy, and R. G. Pearson. 2013.

657 "Ecological Divergence and Speciation between Lemur (Eulemur) Sister Species in

658 Madagascar." *Journal of Evolutionary Biology* 26 (8): 1790–1801.

659 Bouckaert, Remco, Joseph Heled, Denise Kühnert, Tim Vaughan, Chieh-Hsi Wu, Dong

660 Xie, Marc A. Suchard, Andrew Rambaut, and Alexei J. Drummond. 2014. "BEAST

661 2: A Software Platform for Bayesian Evolutionary Analysis." *PLoS Computational*

662 *Biology* 10 (4): e1003537.

663 Burbrink, Frank T., and Marcelo Gehara. 2018. "The Biogeography of Deep Time

664 Phylogenetic Reticulation." *Systematic Biology* 67 (5): 743–44.

665 Cadena, C. D., K. H. Kozak, J. P. Gomez, J. L. Parra, C. M. McCain, R. C. K. Bowie, A.

666 C. Carnaval, et al. 2012. "Latitude, Elevational Climatic Zonation and Speciation in

667 New World Vertebrates." *Proceedings of the Royal Society B: Biological Sciences*.

668 https://doi.org/10.1098/rspb.2011.0720.

669 Cheng, Hai, Ashish Sinha, Francisco W. Cruz, Xianfeng Wang, R. Lawrence Edwards,

670 Fernando M. d'Horta, Camila C. Ribas, Mathias Vuille, Lowell D. Stott, and Augusto

671 S. Auler. 2013. "Climate Change Patterns in Amazonia and Biodiversity." *Nature*

672 *Communications* 4: 1411.

673 Claramunt, S., and J. Cracraft. 2015. "A New Time Tree Reveals Earth History's Imprint

674 on the Evolution of Modern Birds." *Science Advances* 1: e1501005.

675 Coyne, Jerry A., and H. Allen Orr. 2004. "Speciation. Sunderland, MA." Sinauer

676 Associates, Inc.

677 Csilléry, Katalin, Olivier François, and Michael G. B. Blum. 2012. "Abc: An R Package

678 for Approximate Bayesian Computation (ABC)." *Methods in Ecology and Evolution /*

679 *British Ecological Society* 3 (3): 475–79.

680 Delmore, Kira E., Sariel Hübner, Nolan C. Kane, Richard Schuster, Rose L. Andrew,

681 Francisco Câmara, Roderic Guigó, and Darren E. Irwin. 2015. "Genomic Analysis

682 of a Migratory Divide Reveals Candidate Genes for Migration and Implicates

683   Selective Sweeps in Generating Islands of Differentiation." *Molecular Ecology* 24
684   (8): 1873–88.

685   De Queiroz, Kevin. 2007. "Species Concepts and Species Delimitation." *Systematic*
686   *Biology* 56 (6): 879–86.

687   Di Cola, Valeria, Olivier Broennimann, Blaise Petitpierre, Frank T. Breiner, Manuela
688   D'Amen, Christophe Randin, Robin Engler, et al. 2017. "Ecospat: An R Package to
689   Support Spatial Analyses and Modeling of Species Niches and Distributions."
690   *Ecography* 40 (6): 774–87.

691   Ehrendorfer, F. 1959. "Differentiation-Hybridization Cycles and Polyploidy in Achillea."
692   *Cold Spring Harbor Symposia on Quantitative Biology* 24: 141–52.

693   Evanno, G., S. Regnaut, and J. Goudet. 2005. "Detecting the Number of Clusters of
694   Individuals Using the Software STRUCTURE: A Simulation Study." *Molecular*
695   *Ecology* 14 (8): 2611–20.

696   Faircloth, B. C. 2016. "PHYLUCE Is a Software Package for the Analysis of Conserved
697   Genomic Loci." *Bioinformatics* 32 (5): 786–88.

698   Fitzpatrick, Benjamin M., J. A. Fordyce, and S. Gavrilets. 2008. "What, If Anything, Is
699   Sympatric Speciation?" *Journal of Evolutionary Biology* 21 (6): 1452–59.

700   Flantua, Suzette G. A., and Henry Hooghiemstra. 2018. "Historical Connectivity and
701   Mountain Biodiversity." *Mountains, Climate and Biodiversity (eds. Hoorn, C. ,*
702   *Perrigo, A. & Antonelli, A.)*, 171–86.

703   Flantua, Suzette G. A., Aaron O'Dea, Renske E. Onstein, Catalina Giraldo, and Henry
704   Hooghiemstra. 2019. "The Flickering Connectivity System of the North Andean
705   Páramos." *Journal of Biogeography* 41: 1227.

706   Gehara, Marcelo, Adrian A. Garda, Fernanda P. Werneck, Eliana F. Oliveira, Emanuel
707   M. da Fonseca, Felipe Camurugi, Felipe de M. Magalhães, et al. 2017. "Estimating
708   Synchronous Demographic Changes across Populations Using hABC and Its
709   Application for a Herpetological Community from Northeastern Brazil." *Molecular*
710   *Ecology* 26 (18): 4756–71.

711   Gehara, Marcelo, Guilherme G. Mazzochinni, and Frank Burbrink. 2019. "PipeMaster:
712   Inferring Population Divergence and Demographic History with Approximate
713   Bayesian Computation and Supervised Machine-Learning in R."

714   *Compphylo.github.io*.

715   https://compphylo.github.io/Oslo2019/PM_files/PipeMaster.pdf.

716   Gompert, Zachariah, James A. Fordyce, Matthew L. Forister, Arthur M. Shapiro, and

717   Chris C. Nice. 2006. "Homoploid Hybrid Speciation in an Extreme Habitat." *Science*

718   314 (5807): 1923–25.

719   Gompert, Zachariah, Lauren K. Lucas, C. Alex Buerkle, Matthew L. Forister, James A.

720   Fordyce, and Chris C. Nice. 2014. "Admixture and the Organization of Genetic

721   Diversity in a Butterfly Species Complex Revealed through Common and Rare

722   Genetic Variants." *Molecular Ecology* 23 (18): 4555–73.

723   Haffer, J. 1969. "Speciation in Amazonian Forest Birds." *Science* 165 (3889): 131–37.

724   Harrison, Richard G. 2012. "The Language of Speciation." *Evolution* 66 (12): 3643–57.

725   Hewitt, G. 2000. "The Genetic Legacy of the Quaternary Ice Ages." *Nature* 405 (6789):

726   907–13.

727   ———. 2004. "Genetic Consequences of Climatic Oscillations in the Quaternary."

728   *Philosophical Transactions of the Royal Society of London. Series B, Biological*

729   *Sciences* 359 (1442): 183–95.

730   He, Ziwen, Xinnian Li, Ming Yang, Xinfeng Wang, Cairong Zhong, Norman C. Duke,

731   Chung-I Wu, and Suhua Shi. 2019. "Speciation with Gene Flow via Cycles of

732   Isolation and Migration: Insights from Multiple Mangrove Taxa." *National Science*

733   *Review* 6 (2): 275–88.

734   Hoskin, Conrad J., Megan Higgie, Keith R. McDonald, and Craig Moritz. 2005.

735   "Reinforcement Drives Rapid Allopatric Speciation." *Nature* 437 (7063): 1353–56.

736   Hubisz, Melissa J., Daniel Falush, Matthew Stephens, and Jonathan K. Pritchard. 2009.

737   "Inferring Weak Population Structure with the Assistance of Sample Group

738   Information." *Molecular Ecology Resources* 9 (5): 1322–32.

739   Hudson, Richard R. 2002. "Generating Samples under a Wright–Fisher Neutral Model

740   of Genetic Variation." *Bioinformatics* 18 (2): 337–38.

741   Huntley, Jerry W., Johanna A. Harvey, Marco Pavia, Giovanni Boano, and Gary

742   Voelker. 2018. "The Systematics and Biogeography of the Bearded Greenbuls

743   (Aves: Criniger) Reveals the Impact of Plio-Pleistocene Forest Fragmentation on

744   Afro-Tropical Avian Diversity." *Zoological Journal of the Linnean Society* 183 (3):

745   672–86.

746 Huntley, Jerry W., Katrina D. Keith, Adrian A. Castellanos, Lukas J. Musher, and Gary
747   Voelker. 2019. "Underestimated and Cryptic Diversification Patterns across Afro-
748   Tropical Lowland Forests." *Journal of Biogeography* 104 (January): 20432.

749 Irwin, Darren E., Borja Milá, David P. L. Toews, Alan Brelsford, Haley L. Kenyon, Alison
750   N. Porter, Christine Grossen, Kira E. Delmore, Miguel Alcaide, and Jessica H.
751   Irwin. 2018. "A Comparison of Genomic Islands of Differentiation across Three
752   Young Avian Species Pairs." *Molecular Ecology* 27 (23): 4839–55.

753 Jakobsson, Mattias, and Noah A. Rosenberg. 2007. "CLUMPP: A Cluster Matching and
754   Permutation Program for Dealing with Label Switching and Multimodality in
755   Analysis of Population Structure." *Bioinformatics* 23 (14): 1801–6.

756 Kiehl, Jeffrey T., and Peter R. Gent. 2004. "The Community Climate System Model,
757   Version 2." *Journal of Climate* 17 (19): 3666–82.

758 Kozak, Kenneth H., and John J. Wiens. 2006. "Does Niche Conservatism Promote
759   Speciation? A Case Study in North American Salamanders." *Evolution* 60 (12):
760   2604–21.

761 Leaché, Adam D., and Matthew K. Fujita. 2010. "Bayesian Species Delimitation in West
762   African Forest Geckos (Hemidactylus Fasciatus)." *Proceedings of the Royal*
763   *Society B: Biological Sciences*, October. https://doi.org/10.1098/rspb.2010.0662.

764 Lerner, Heather R. L., Matthias Meyer, Helen F. James, Michael Hofreiter, and Robert
765   C. Fleischer. 2011. "Multilocus Resolution of Phylogeny and Timescale in the
766   Extant Adaptive Radiation of Hawaiian Honeycreepers." *Current Biology: CB* 21
767   (21): 1838–44.

768 Marques, David A., Joana I. Meier, and Ole Seehausen. 2019. "A Combinatorial View
769   on Speciation and Adaptive Radiation." *Trends in Ecology & Evolution* 34 (6): 531–
770   44.

771 Mayr, Ernst, and Robert J. O'Hara. 1986. "THE BIOGEOGRAPHIC EVIDENCE
772   SUPPORTING THE PLEISTOCENE FOREST REFUGE HYPOTHESIS." *Evolution*
773   40 (1): 55–67.

774 Mirarab, S., R. Reaz, Md S. Bayzid, T. Zimmermann, M. S. Swenson, and T. Warnow.
775   2014. "ASTRAL: Genome-Scale Coalescent-Based Species Tree Estimation."

776   *Bioinformatics* 30 (17): i541–48.

777 Morey, Richard D., Jeffrey N. Rouder, and T. Jamil. 2015. "BayesFactor: Computation
778   of Bayes Factors for Common Designs." *R Package Version 0. 9* 9: 2014.

779 Muscarella, Robert, Peter J. Galante, Mariano Soley-Guardia, Robert A. Boria, Jamie
780   M. Kass, María Uriarte, and Robert P. Anderson. 2014. "ENMeval: An R Package
781   for Conducting Spatially Independent Evaluations and Estimating Optimal Model
782   Complexity for Maxent Ecological Niche Models." *Methods in Ecology and*
783   *Evolution* 5 (11): 1198–1205.

784 Musher, Lukas J., and Joel Cracraft. 2018. "Phylogenomics and Species Delimitation of
785   a Complex Radiation of Neotropical Suboscine Birds (Pachyramphus)." *Molecular*
786   *Phylogenetics and Evolution* 118 (January): 204–21.

787 Musher, Lukas J., Mateus Ferreira, Anya L. Auerbach, Jessica McKay, and Joel
788   Cracraft. 2019. "Why Is Amazonia a 'source'of Biodiversity? Climate-Mediated
789   Dispersal and Synchronous Speciation across the Andes in an Avian Group
790   (Tityrinae)." *Proceedings of the Royal Society B* 286 (1900): 20182343.

791 Norris, Laura C., Bradley J. Main, Yoosook Lee, Travis C. Collier, Abdrahamane
792   Fofana, Anthony J. Cornel, and Gregory C. Lanzaro. 2015. "Adaptive Introgression
793   in an African Malaria Mosquito Coincident with the Increased Usage of Insecticide-
794   Treated Bed Nets." *Proceedings of the National Academy of Sciences* 112 (3):
795   815–20.

796 Nosil, Patrik. 2012. *Ecological Speciation*. Oxford University Press.

797 Oswald, Jessica A., Michael G. Harvey, Rosalind C. Remsen, Depaul U. Foxworth,
798   Donna L. Dittmann, Steven W. Cardiff, and Robb T. Brumfield. 2019. "Evolutionary
799   Dynamics of Hybridization and Introgression Following the Recent Colonization of
800   Glossy Ibis (Aves: Plegadis Falcinellus) into the New World." *Molecular Ecology*
801   28(7), 1675–1691. https://doi.org/10.1111/mec.15008.

802 Oswald, Jessica A., Isaac Overcast, William M. Mauck 3rd, Michael J. Andersen, and
803   Brian Tilston Smith. 2017. "Isolation with Asymmetric Gene Flow during the
804   Nonsynchronous Divergence of Dry Forest Birds." *Molecular Ecology* 26 (5): 1386–
805   1400.

806 Otto-Bliesner, Bette L., Shawn J. Marshall, Jonathan T. Overpeck, Gifford H. Miller, and

807   Aixue Hu. 2006. "Simulating Arctic Climate Warmth and Icefield Retreat in the Last
808   Interglaciation." *Science* 311 (5768): 1751–53.

809 Parker, T. A., D. F. Stotz, J. W. Fitzpatrick, and D. K. Moskovits. 1996. "Database A:
810   Zoogeographic and Ecological Attributes of Bird Species Breeding in the
811   Neotropics." In *Neotropical Birds: Ecology and Conservation*, edited by Douglas F.
812   Stotz, John W. Fitzpatrick, Theodore A. Parker III, Debra K. Moskovits, 132–291.
813   University of Chicago Press, Chicago.

814 Pavlidis, Pavlos, Stefan Laurent, and Wolfgang Stephan. 2010. "msABC: A Modification
815   of Hudson's Ms to Facilitate Multi-Locus ABC Analysis." *Molecular Ecology*
816   *Resources* 10 (4): 723–27.

817 Pfeifer, Bastian, Ulrich Wittelsbürger, Sebastian E. Ramos-Onsins, and Martin J.
818   Lercher. 2014. "PopGenome: An Efficient Swiss Army Knife for Population
819   Genomic Analyses in R." *Molecular Biology and Evolution* 31 (7): 1929–36.

820 Phillips, Steven J., Robert P. Anderson, Miroslav Dudík, Robert E. Schapire, and Mary
821   E. Blair. 2017. "Opening the Black Box: An Open-Source Release of Maxent."
822   *Ecography* 40 (7): 887–93.

823 Phillips, Steven J., Robert P. Anderson, and Robert E. Schapire. 2006. "Maximum
824   Entropy Modeling of Species Geographic Distributions." *Ecological Modelling* 190
825   (3): 231–59.

826 Pickrell, Joseph K., and Jonathan K. Pritchard. 2012. "Inference of Population Splits and
827   Mixtures from Genome-Wide Allele Frequency Data." *PLoS Genetics* 8 (11):
828   e1002967.

829 Potter, Sally, Alexander T. Xue, Jason G. Bragg, Dan F. Rosauer, Emily J. Roycroft,
830   and Craig Moritz. 2018. "Pleistocene Climatic Changes Drive Diversification across
831   a Tropical Savanna." *Molecular Ecology*. https://doi.org/10.1111/mec.14441.

832 Prates, Ivan, Danielle Rivera, Miguel T. Rodrigues, and Ana C. Carnaval. 2016. "A Mid-
833   Pleistocene Rainforest Corridor Enabled Synchronous Invasions of the Atlantic
834   Forest by Amazonian Anole Lizards." *Molecular Ecology* 25 (20): 5174–86.

835 Provine, William B. 2004. "Ernst Mayr: Genetics and Speciation." *Genetics* 167 (3):
836   1041–46.

837 Provost, Kaiya L., William M. Mauck 3rd, and Brian Tilston Smith. 2018. "Genomic

838 Divergence in Allopatric Northern Cardinals of the North American Warm Deserts Is
839 Linked to Behavioral Differentiation." *Ecology and Evolution* 8 (24): 12456–78.

840 Racimo, Fernando, Sriram Sankararaman, Rasmus Nielsen, and Emilia Huerta-
841 Sánchez. 2015. "Evidence for Archaic Adaptive Introgression in Humans." *Nature*
842 *Reviews. Genetics* 16 (6): 359–71.

843 Rannala, Bruce, and Ziheng Yang. 2013. "Improved Reversible Jump Algorithms for
844 Bayesian Species Delimitation." *Genetics* 194 (1): 245–53.

845 Rattenbury, J. A. 1962. "Cyclic Hybridization as a Survival Mechanism in the New
846 Zealand Forest Flora." *Evolution* 16 (3): 348–63.

847 Robinson, W. Douglas, Jeffrey D. Brawn, and Scott K. Robinson. 2000. "Forest Birds
848 Community Structure in Central Panama: Influence of Spatial and Ecological Scale
849 and Biogeography." *Ecological Monographs* 70 (2): 209–35.

850 Rull, Valentí, and Sandra Nogué. 2007. "Potential Migration Routes and Barriers for
851 Vascular Plants of the Neotropical Guyana Highlands during the Quaternary."
852 *Journal of Biogeography* 34 (8): 1327–41.

853 Sanmartín, Isabel, Paul Van Der Mark, and Fredrik Ronquist. 2008. "Inferring Dispersal:
854 A Bayesian Approach to Phylogeny-Based Island Biogeography, with Special
855 Reference to the Canary Islands." *Journal of Biogeography* 35 (3): 428–49.

856 Schilthuizen, M. 2000. "Dualism and Conflicts in Understanding Speciation." *BioEssays*
857 22 (12): 1134–41.

858 Schluter, D. 2001. "Ecology and the Origin of Species." *Trends in Ecology & Evolution*
859 16 (7): 372–80.

860 Schoener, Thomas W. 1968. "The Anolis Lizards of Bimini: Resource Partitioning in a
861 Complex Fauna." *Ecology* 49 (4): 704–26.

862 Slatkin, M. 1987. "Gene Flow and the Geographic Structure of Natural Populations."
863 *Science* 236 (4803): 787–92.

864 Smith, Brian Tilston, Amei Amei, and John Klicka. 2012. "Evaluating the Role of
865 Contracting and Expanding Rainforest in Initiating Cycles of Speciation across the
866 Isthmus of Panama." *Proceedings. Biological Sciences / The Royal Society* 279
867 (1742): 3520–26.

868 Stamatakis, Alexandros. 2014. "RAxML Version 8: A Tool for Phylogenetic Analysis and

869   Post-Analysis of Large Phylogenies." *Bioinformatics* 30 (9): 1312–13.

870 Thom, Gregory, Fabio Raposo Do Amaral, Michael J. Hickerson, Alexandre Aleixo,
871   Lucas E. Araujo-Silva, Camila C. Ribas, Erik Choueri, and Cristina Y. Miyaki. 2018.
872   "Phenotypic and Genetic Structure Support Gene Flow Generating Gene Tree
873   Discordances in an Amazonian Floodplain Endemic Species." *Systematic Biology*
874   67 (4): 700–718.

875 Tingley, Morgan W., William B. Monahan, Steven R. Beissinger, and Craig Moritz. 2009.
876   "Birds Track Their Grinnellian Niche through a Century of Climate Change."
877   *Proceedings of the National Academy of Sciences* 106: 19637–43.

878 Trujillo-Arias, Natalia, Luciano Calderón, Fabricio R. Santos, Cristina Y. Miyaki,
879   Alexandre Aleixo, Christopher C. Witt, Pablo L. Tubaro, and Gustavo S. Cabanne.
880   2018. "Forest Corridors between the Central Andes and the Southern Atlantic
881   Forest Enabled Dispersal and Peripatric Diversification without Niche Divergence in
882   a Passerine." *Molecular Phylogenetics and Evolution* 128 (November): 221–32.

883 Vrba, Elisabeth S. 1992. "Mammals as a Key to Evolutionary Theory." *Journal of*
884   *Mammalogy* 73 (1): 1–28.

885 Vuilleumier, B. S. 1971. "Pleistocene Changes in the Fauna and Flora of South
886   America." *Science* 173 (3999): 771–80.

887 Webb, S. David. 1991. "Ecogeography and the Great American Interchange."
888   *Paleobiology* 17 (3): 266–80.

889 Webster, J. Dan. 1963. "A Revision of the Rose-Throated Becard." *The Condor* 65 (5):
890   383–99.

891 Weeks, Brian C., and Santiago Claramunt. 2014. "Dispersal Has Inhibited Avian
892   Diversification in Australasian Archipelagoes." *Proceedings. Biological Sciences /*
893   *The Royal Society* 281 (1791): 20141257.

894 Weir, Jason T., Maya S. Faccio, Paola Pulido-Santacruz, Alfredo O. Barrera-Guzmán,
895   and Alexandre Aleixo. 2015. "Hybridization in Headwater Regions, and the Role of
896   Rivers as Drivers of Speciation in Amazonian Birds." *Evolution* 69 (7): 1823–34.

897 Wiens, John J., and Catherine H. Graham. 2005. "Niche Conservatism: Integrating
898   Evolution, Ecology, and Conservation Biology." *Annual Review of Ecology,*
899   *Evolution, and Systematics* 36 (1): 519–39.

900 Winger, Benjamin M. 2017. "Consequences of Divergence and Introgression for

901   Speciation in Andean Cloud Forest Birds." *Evolution* 71 (7): 1815–31.

902 Yu, Yun, and Luay Nakhleh. 2015. "A Maximum Pseudo-Likelihood Approach for

903   Phylogenetic Networks." *BMC Genomics* 16: S10.

904 Zarza, Eugenia, Brant C. Faircloth, Whitney L. E. Tsai, Robert W. Bryson Jr, John

905   Klicka, and John E. McCormack. 2016. "Hidden Histories of Gene Flow in Highland

906   Birds Revealed with Genomic Markers." *Molecular Ecology* 25 (20): 5144–57.

907

908 **BIOSKETCH:**

909 LJM is a PhD candidate at the American Museum of Natural History who combines

910 phylogenomics, biogeography, and population genomics to test questions about the

911 origins and assembly of avian communities in the Neotropics. He is particularly

912 interested in the Amazon, as it is the most diverse terrestrial biome on the planet. Using

913 high-throughput sequencing technologies and specimen-based research, LJM asks

914 questions about how landscape evolution –such as that driven by climate change or

915 watershed dynamics– have shaped patterns of population genetic and taxonomic

916 diversity through facilitating dispersal, isolation, and gene flow.

N

Localities

- **P. aglaiae**
- **P. sumichrasti**
- **P. canescens**
- **P. homochrous**

500   0   500   1000   1500   2000 km